DOCUMENT RESUME

ED 320 135 CS 010 134

AUTHOR Goetz, Ernest T.; And Others
TITLE Getting a Reading on Reader Response: Relationships between Ratings, Recall, and Imagery Reports.
PUB DATE Jan 90
NOTE 27p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Austin, TX, January 25-27, 1990).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Data Interpretation; Emotional Response; Higher Education; *Imagery; *Reader Response; Reader Text Relationship; Reading Comprehension; *Reading Processes; Reading Research; *Reading Skills; Reading Strategies; *Recall (Psychology)
IDENTIFIERS Paivio (A)

ABSTRACT

A study reanalyzed data from two previous studies to investigate (1) validity of imagery ratings and reports as measures of reader response to a story, and (2) similarity of final regression models for recall and imagery reports. Imagery and affect ratings of story paragraphs were found to predict both imagery reports and recall of a second, independent group of readers, lending convergent construct validity to these measures. Hierarchical regression analyses of within-subjects variance ensured that rating effects could not be attributed to surface-level text factors. Comparison of final regression models for imagery and recall revealed differences that appeared to support Paivio's dual code hypothesis. (Four tables of data are included, and 22 references are attached.) (Author/MG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED320135

CS

Getting a Reading on Reader Response: Relationships between Ratings, Recall, and Imagery Reports

Ernest T. Goetz

Mark Sadoski

Arturo Olivarez, Jr.

Texas A&M University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

ERNEST T. GOETZ

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CS010134

Running Head: Reader Response

BEST COPY AVAILABLE

Abstract

Data from two previous investigations (Sadoski & Goetz, 1985; Sadoski, Goetz, Olivarez, Lee, & Roberts, in press) were reanalyzed to investigate: (1) validity of imagery ratings and reports as measures of reader response to a story, (2) similarity of final regression models for recall and imagery reports. Imagery and affect ratings of story paragraphs were found to predict both imagery reports and recall of a second, independent group of readers, lending convergent construct validity to these measures. Hierarchical regression analyses of within-subjects variance (cf. Anderson, Mason, & Shirey, 1984; Herman, Anderson, Pearson, & Nagey, 1987) ensured that rating effects could not be attributed to surface-level text factors. Comparison of final regression models for imagery and recall revealed differences that appeared to support Paivio's (1971, 1986) dual code hypothesis.

The view of text comprehension and memory that has emerged in the last decade is that of an interaction between reader and text in which the reader actively constructs a representation of the objects and events depicted in the text. Thus, text processing is said to be constructive, generative, or elaborative. For stories, the reader may construct images of characters, scenes, and events that breathe life into the text. The reader may experience emotional responses, empathizing with story characters or reacting to story events.

Paradoxically, the emphasis on the active, constructive role of the reader has not brought wide attention to the reader's imagery or affective experiences. Most text research has focused on comprehension and recall, and most imagery research has focused on tasks other than reading. Readers' experiences of imagery and affect remain underrepresented in research on reader response.

The paucity of research on readers' imagery and affective experiences may reflect concerns about the inaccessibility and subjectivity of such experiences. Suspicions about the reliability and validity of imagery and affect reports may have biased researchers to neglect such data. More generally, concerns have been voiced

regarding the validity of all verbal reports of mental processes by Nisbett and Wilson (1977) and Evans (1980). These criticisms have been answered by Ericsson and Simon (1980, 1984; see also Morris, 1981a, 1981b) who provided a theoretical framework (i.e., human information processing) and procedural guidelines for collecting valid verbal report data. One of the areas in which verbal report data have been shown reliable and informative is in recent studies of imagery and affect in reading. Likert-type ratings of readers' imagery and emotional responses to story paragraphs have proven highly reliable, with alpha reliability coefficients generally above .9 (Goetz, Sadoski, Stowe, Fetsco, & Kemp, 1987; Katz, Paivio, & Marschark, 1985; Sadoski, Goetz, & Kangiser, 1988). Retrospective free reports of images experienced during reading (Sadoski, Goetz, Olivarez, Lee, & Roberts, in press) is another methodology that has proven useful. Proportion of interrater agreement for coding decisions of imagery reports has been generally above .8, suggesting acceptable reliability for such data.

The present investigation explored the relationship between these two methodologies by examining the interrelationships between data collected in two previous

studies (Sadoski & Goetz, 1985; Sadoski et al., in press). Specifically, this investigation was intended to determine whether: (1) imagery, affect, and importance ratings of one group of readers predicted the retrospective free imagery reports and free recall of a second set of readers, and (2) whether the same variables predicted imagery reports and recall. If recall and imagery reports can be shown to be predicted by the ratings of an independent set of readers, after controlling for surface level text factors (e.g., syllables per word, words per sentence), then convergent construct validity (Campbell & Fiske, 1959) for these measures will be demonstrated. If different variables are found to predict imagery reports and recall, then that would argue for the need to separately address imagery and verbal recall, and suggest support for Paivio's (1971, 1986) dual code hypothesis.

Method

Summary of Procedures and Description of Data

Imagery, Affect, and Importance Ratings. Sadoski and Goetz (1985) explored reader responses by having readers rate story paragraphs. The story, "First Kill," was a 2,100 word excerpt from an adolescent novel (Annixter & Annixter, 1958). "First Kill" recounts the

rite of passage to tribal hunter of a Sioux Indian youth, who proves himself by single-handedly killing a buffalo on his first hunt. The story describes the youth's thoughts and feelings as well as the action of the hunt.

Three groups of 15 undergraduate education students each read the story and then rated story paragraphs on one of the following five-point Likert-type scales: imagery ("no imagery" to "vivid imagery"), affect ("no emotional reaction" to "very strong emotional reaction"), or importance ("incidental information" to "key occurrences critical to the plot and theme"). All students read a typed copy of the story (no illustrations), and then were presented with a second copy of the story in which the appropriate rating scale appeared next to story paragraphs. Students rated each of 28 paragraphs (seven very brief paragraphs with one or two sentences were not rated) on the scale provided.

The data used in the present investigation were the mean ratings for each paragraph. The imagery, affect, and importance means for the 28 paragraphs were calculated by averaging across the 15 students in each condition.

Recall and Imagery Reports. Sadoski et al. (in press) investigated reports of spontaneous mental imagery

and recall for the same story. Seventy-two community college students read "First Kill" under instructions to read for pleasure, to look for the author's theme, or to find and mark typographical errors that had been inserted in the text for that group. Immediately after reading the story, all students produced written imagery reports and free recalls in counterbalanced order. For the imagery reports, students were asked to number and briefly describe each image they recalled from reading the story. Although it had been anticipated that the instructions for reading the story would lead to different levels of processing of the text and therefore influence imagery and recall, neither reading instructions nor task order (recall or imagery reports first) produced significant effects.

The data analyzed in the present investigation were the immediate recalls and imagery reports. The recall protocols were scored with the T-unit (a main clause with all of its subordinate clauses; Hunt, 1965) as the unit of analysis. Each recall statement was first classified as: (a) identifiable with a single story T-unit, (b) a synthesis of information from several T-units, or (c) a reader importation. Recall identified with a T-unit was classified as a reasonably accurate representation of the

gist of the text unit or as having major distortions or omissions. For the present investigation, recall was collapsed across T-units for each of the 28 paragraphs rated in the previous study (Sadoski & Goetz, 1985). Total recall, regardless of accuracy, and gist recall were analyzed.

Imagery reports could only be reliably associated with paragraphs, and were classified as: (a) directly related to a single story paragraph, (b) syntheses of information from several story paragraphs, or (c) reader importations. Each imagery report was then coded for modality (e.g., visual, auditory, affective) using a conservative scoring criterion that resulted in more than half of the imagery reports being classified as nondeterminate for modality. For the present study, total images and visual images identified with the 28 rated paragraphs were analyzed.

Analyses

The full subjects X paragraphs data matrices (72 x 28 = 2016 observations) for total and gist recall, and total and visual image reports from Sadoski et al. (in press) were analyzed following Anderson's (e.g., Anderson, Mason, & Shirey, 1984; Herman, Anderson, Pearson, & Nagey, 1987) approach to within-subjects

hierarchical regression analyses. For both imagery report and recall measures, all between-subjects variance was removed, and the remaining, within-subjects variance was analyzed in the following sequence of stages: (1) surface level text characteristics of each of the 28 paragraphs (the number of T-units, words, propositions, and syllables in the paragraph; average syllables per word, words per sentence, and propositions per sentence for the paragraph), (2) quadratics of text characteristics (to account for possible nonlinear relationships), (3) average imagery, affect, and importance ratings from a different sample of subjects (Sadoski & Goetz, 1985), (4) quadratics of these ratings, (5) interactions of between-subjects factors (age, sex, native language -- English or other, Stanford Diagnostic Reading Test [1977] scores, experimental instructions, and order of imagery reports and recall) with the ratings.

The set of variables entered in each stage was evaluated by step-wise regression procedures. Because of the large number of observations and effects being tested, the alpha level for inclusion as a predictor variable was set at $p < .01$ (critical value for 1, 1944 degrees of freedom = 6.68). Variables found significant

at a given stage of the analyses were retained in the model statement for all subsequent stages in the order they originally entered the predictor model, thus providing overall hierarchical regression analyses.

Results

Recall

Table 1 summarizes the analyses of total recall ($\bar{x}$ = 1.71, *SD* = 0.94). Between-subjects variance amounted to 13.33% of total variance. The R^2 values in Table 1 (and subsequent tables) and the following text refer to within-subjects variance accounted for at the time each variable entered the model. Since recall was scored in T-units, the number of T-units in the paragraph was the first variable entered in stage 1, and it accounted for 17.65% of the variance. Four other text factors (propositions per paragraph, words per sentence, propositions per sentence, words per paragraph) accounted for an additional 8.59%. The quadratic of propositions per sentence was the only variable to enter in stage 2.

Insert Table 1 about here

After all of the variance associated with surface-level text characteristics had been removed, the relationship between the ratings of one group of readers

and the recall of an independent group were evaluated in stage 3. All three ratings proved highly significant predictors ($p < .001$). The three ratings accounted for 3.44% of the variance, with the imagery rating making the largest contribution, followed closely by the importance rating. Stage 4 revealed an additional contribution for the quadratic of the importance rating. The final stage revealed significant interactions for age and the affect rating and native language and the imagery rating. Total variance added in each of the last two stages, however, was less than 1%.

The regression analyses for recall that represented the gist of the text ($\bar{x} = 0.90$, $SD = 0.92$) are summarized in Table 2. Between-subjects variance was 16.29% of the total. Number of T-units in the paragraph was again entered first in stage 1, accounting for 17.21% of the within-subjects variance. Number of words and number of sentences combined for an additional 3.82%. In stage 2, five quadratics added 3.89%. Stage 3 of the analyses revealed that all three ratings from the separate group of subjects predicted gist recall apart from surface level text factors. The ratings predicted an additional 3.65%, with affect making the largest contribution, followed by imagery. The quadratics of imagery and

importance ratings added 2.48% in stage 4. In stage 5, three interactions proved significant, accounting for a total of 1.13% variance.

Insert Table 2 about here

Imagery

The analyses of the total number of imagery reports identified with the 28 paragraphs analyzed ($\bar{x}$ = 0.40, SD = 0.55) are summarized in Table 3. Between-subjects variance constituted 16.27 % of the total. Although imagery was not scored by T-units, this variable was entered first in stage 1 to maintain a parallel to the recall analyses, and accounted for 5.85% of the within-subjects variance. Words per sentence and number of syllables in the paragraph added 5.59%. In stage 2, the quadratic of T-units entered the model.

Insert Table 3 about here

Stage 3, which evaluated the relationship between ratings and imagery reports of independent groups of readers, revealed highly significant effects for imagery and affect ratings, but none for importance ratings, once surface level text characteristics had been accounted

for. Imagery and affect ratings combined for 3.87% of the variance, with imagery ratings making up most of that sum. In stage 4, the quadratic of the importance rating was added, and in stage 5, the interactions of native language and order with affect ratings combined for an additional 0.73%.

Table 4 summarizes the analyses of the imagery reports classified as visual in modality ($\bar{X}$ = 0.15, *SD* = 0.36). To maintain the parallel with previous analyses, T-units were entered first and retained although they did not reach significance. Perhaps due to floor effects, only three variables proved significant predictors: number of words in the paragraph in stage 1, and the imagery and affect ratings of separate subjects in stage 3.

Insert Table 4 about here

Imagery Reports and Recall

An ancillary analysis investigated the relationship between imagery reports and recall. Treating total recall as the dependent variable and controlling for surface-level text factors (Table 1, Stages 1 & 2), total number of reported images was highly significant, *F* =

55.17, $p < .0001$, $R^2 = .0195$. Reported images and recall were positively related (B value = 0.3182).

Discussion

The results of these analyses provide evidence of the convergent construct validity of imagery reports and of imagery and affective response ratings as indices of the reader's experience. Imagery and affect ratings of story paragraphs from one group of readers were found to be highly significant predictors of imagery reports of an independent group of readers. Imagery and affect ratings (along with importance ratings) also predicted story recall of other readers. As might be expected, the strongest relationship found was between imagery ratings and total reported images. These relationships could not be accounted for in terms of surface-level text factors, since the plan of the analyses ensured that any significant variance attributable to such factors was removed prior to the evaluation of rating effects.

The results of this study are both methodologically and theoretically important. Methodologically, the study is important because it establishes that the relationship between imagery and affective responses during reading (i.e., the ratings) and subsequent story recall is not an artifact of having the same subjects rate the story and

then recall it. Since the ratings and the recalls and imagery reports were obtained from separate groups of people, this possible confound was eliminated. Further, this correlational finding is consistent with experimental studies in which imagery has been manipulated by instructions and presumably played a causal role in recall (Paivio, 1971, 1986).

The study is theoretically important because different patterns of predictors emerged from the analyses of recall and imagery reports. Comparison of the final regression models for total recall and total images reported reveals two differences. First, surface-level text factors accounted for more variance in predicting total recall than in predicting total imagery reports (26.24% vs. 11.44% for stage 1). A test of this difference employing Fisher's z-transformation of correlation coefficients confirmed its significance, $z=6.79$, $p<.001$. Second, importance ratings proved to be significant predictors of recall but not of imagery reports. This pattern of results, which held for gist recall and visual imagery reports as well, is generally consistent with Paivio's (1971, 1986) dual code hypothesis, which holds that verbal and imagery codes constitute distinct but interrelated modes of

representation in memory (as opposed to a common, amodal memory code). Ratings of the importance of paragraphs to the plot hierarchy and theme of the story and measures of text factors such as syllables, words, and propositions, being primarily linguistic entities related to the verbal or propositional code, were found to be more closely related to verbal recall than to imagery reports, while imagery ratings were most strongly related to imagery reports. It might be argued that floor effects and the slightly lower reliability of imagery reports suppressed observed relationships between predictors and imagery reports, but the fact that the imagery rating accounted for as much variance in the imagery reports as all three ratings combined did for recall mitigates against that argument. For visual images, however, floor effects were a serious limitation.

Overall, the final regression models may appear to account for only a modest amount of variance. This can be attributed to analyses that treated a single subject's recall or imagery report for a single paragraph as the unit of analysis, rather than aggregating data across subjects or across the text, as is typically done. Aggregation of results causes large increases in the amount of variation apparently explained, but ignores

systematic individual differences between subjects, and interactions between individual difference variables and experimental variables (Anderson et al., 1984; Herman et al., 1987). The present approach to analysis provides a more sensitive, fine-grained, and stable test of effects, that: (a) removed all between-subjects variance, (b) entailed a large number of observations, and (c) maintained information that would be lost through aggregation. In the present analyses, interactions of between-subjects factors and ratings were observed, but their contributions were small relative to the main effects of the ratings, and no pattern of interactions emerged across the analyses of recall and imagery.

The present study employed a single text, but a relationship between imagery and recall has been found previously in investigations of third and fourth graders reading stories from basal readers (Sadoski, 1983, 1985). Further, a recent investigation in which college readers read and rated the paragraphs of three articles from popular magazines (Sadoski & Quast, 1988) revealed that imagery and affect ratings remained significant predictors of recall after a two-week delay.

The present results support the contention that imagery and affect are important aspects of the reader's

experience with a story, and that these aspects of the reader's response can be studied empirically with reliability and validity. Although they are intimately intertwined with verbal comprehension and recall, they are worthy of study in their own right. Imagery and emotional response enrich the reader's experience. The study of these imaginative processes may enrich our understanding of that experience.

References

Anderson, R. C., Mason, J. M., & Shirey, L. (1984). The reading group: An experimental investigation of a labyrinth. Reading Research Quarterly, 20, 6-38.

Annixter, J., & Annixter, P. (1958). Buffalo chief. New York: Holiday House.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-185.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87, 215-251.

Ericsson, K. A., & Simon, H. A. (1984). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.

Evans, J., St. B. T. (1980). Introspective reports of cognitive processes: A critique. Bulletin of the British Psychological Society, 33, 32-33.

Goetz, E. T., Sadoski, M., Stowe, M. L., Kemp, S. G., & Fetsco, T. G. (1987, December). Imagery and emotional response in reading: On-line measures and retrospective reports. Paper presented at the meeting of the National Reading Conference, St. Petersburg Beach, FL.

Herman, P. A., Anderson, R. C., Pearson, P. D., & Nagey, W. E. (1987). Incidental acquisition of word meanings from expositions with varied text features. Reading Research Quarterly, 22, 263-284.

Hunt, K. W. (1965). Grammatical structures written at three grade levels. NTCE Research Report No. 3. Champaign, IL: National Council of Teachers of English.

Katz, A. N., Paivio, A., & Marschark, M. (1985). Poetic comparisons: Psychological dimensions of metaphoric processing. Journal of Psycholinguistic Research, 14, 365-383.

Morris, P. E. (1981a). Why Evans is wrong in criticizing introspective reports of subjects strategies. British Journal of Psychology, 72, 465-468.

Morris, P. E. (1981b). The cognitive psychology of self reports. In C. Antaki (Ed.), The psychology of ordinary explanations of social behavior. London and New York: Academic Press.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 59, 421-430.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart, & Winston. (Reprinted by Lawrence Erlbaum, 1979).

Paivio, A. (1986). Mental representations: A dual coding approach. New York: Oxford University Press.

Sadoski, M. (1983). An exploratory study of the relationships between reported imagery and the comprehension and recall of a story. Reading Research Quarterly, 19, 110-123.

Sadoski, M. (1985). The natural use of imagery in story comprehension and recall: Replication and extension. Reading Research Quarterly, 20, 658-667.

Sadoski, M., & Goetz, E. T. (1985). Relationships between affect, imagery, and importance ratings for segments of a story. In J. A. Niles & R. Lalik (Eds.), Issues in Literacy: A research perspective (pp. 180-185). Thirty-fourth Yearbook of the National Reading Conference. Washington, DC: National Reading Conference.

Sadoski, M., Goetz, E. T., & Kangizer, S. M. (1988). Imagination in story response: Relationships between imagery, affect, and structural importance. Reading Research Quarterly, 23, 320-336.

Sadoski, M., Goetz, E. T., Olivarez, A., Jr., Lee, S., & Roberts, N. M. (in press). Imagination in story reading: The role of imagery, verbal recall, story analysis, and processing levels. Journal of Reading Behavior.

Sadoski, M., & Quast, Z. (1988, November). Relationships between imagery, affect, and structural salience in journalistic text. Paper presented at the meeting of the National Reading Conference, Tucson, AZ.

Stanford diagnostic reading test (1977). Blue level. Karlson, B., Madden, R., & Gardner, E. F. New York: Harcourt, Brace, Jovanovich.

TABLE 1

Final Hierarchical Regression Model for Total Recall

Stage	Variable	b	% Variance	F
1	T-units/Paragraph	0.19	17.65	431.61
	Proposition/Paragraph	0.13	4.92	127.85
	Words/Sentence	0.22	1.31	34.54
	Proposition/Sentence	-0.68	0.33	8.88
	Words/Paragraph	-0.06	2.03	55.43
2	Quadratic of:			
	Proposition/Sentence	0.02	0.81	22.31
3	Affect Rating	0.14	0.40	11.02
	Importance Rating	-0.41	1.42	39.99
	Imagery Rating	0.67	1.63	46.90
4	Quadratic of:			
	Importance Rating	0.20	0.48	13.86
5	Age X Affect Rating	-0.04	0.46	12.60
	Language X Imagery Rating	-0.33	0.40	10.90

TABLE 2

Final Hierarchical Regression Model for Gist Recall

Stage	Variable	b	% Variance	F
1	T-units/Paragraph	0.26	17.21	418.80
	Words/Paragraph	-0.01	2.89	72.91
	Sentence/Paragraph	-0.03	0.92	23.42
2	Quadratic of:			
	Syllables/Paragraph	0.00	0.55	14.08
	Proposition/Paragraph	0.00	0.70	18.01
	Proposition/Sentence	-0.02	1.09	28.60
	Syllables/Word	0.91	1.12	29.86
	Words/Sentence	0.01	0.43	11.61
3	Imagery Rating	1.76	1.12	30.28
	Importance Rating	-0.14	0.37	10.01
	Affect Rating	-0.39	2.16	60.59
4	Quadratic of:			
	Imagery Rating	-0.42	1.68	48.25
	Importance Rating	0.27	0.80	23.24
5	Age X Importance Rating	-0.03	0.43	11.83
	SDRT[a] X Imagery Rating	0.02	0.39	10.89
	Language X Imagery Rating	-0.27	0.31	9.60

Note. [a]Stanford Diagnostic Reading Test.

TABLE 3

Final Hierarchical Regression Model for Total Imagery

Stage	Variable	b	% Variance	F
1	T-unit/Paragraph	-0.03	5.85	125.10
	Words/Sentence	0.01	4.81	108.47
	Syllables/Paragraph	0.00	0.78	17.68
2	Quadratic of:			
	T-units/Paragraph	0.00	0.30	6.74
3	Imagery Rating	0.22	3.39	80.29
	Affect Rating	-0.08	0.48	11.45
4	Quadratic of:			
	Importance Rating	0.02	0.30	7.17
5	Language X Affect Rating	-0.18	0.39	8.78
	Order X Affect Rating	0.10	0.34	7.61

TABLE 4

Final Hierarchical Regression Model for Visual Imagery

Stage	Variable	b	% Variance	F
1	T-units/Paragraph	-0.01	0.16	3.16
	Words/Paragraph	0.00	1.57	32.14
3	Imagery Rating	0.05	0.66	13.70
	Affect Rating	-0.04	0.55	11.44